

Local semicircle law at the spectral edge for Gaussian β -ensembles

Percy Wong

November 8, 2011

Abstract

We study the local semicircle law for Gaussian β -ensembles at the edge of the spectrum. We prove that at the almost optimal level of $n^{-2/3+\epsilon}$, the local semicircle law holds for all $\beta \geq 1$ at the edge. The proof of the main theorem relies on the calculation of the moments of the tridiagonal model of Gaussian β -ensembles up to the p_n -moment where $p_n = O(n^{2/3-\epsilon})$. The result is the analogous to the result of Sinai and Soshnikov [14] for Wigner matrices, but the combinatorics involved in the calculations are different.

1 Introduction and Summary

Our goal in this paper is to prove a local semicircle law at the level of $O(n^{2/3-\epsilon})$ near the spectral edges of the Gaussian β -ensembles. The corresponding problem for Wigner matrices at the edge has been established by Sinai and Soshnikov [14] and in the bulk by Erdős, Schlein and Yau [7], [8].

Definition 1. *Let $\beta \geq 1$, the Gaussian β -ensemble is an ensemble in \mathbb{R} that have the following probability density:*

$$df(\lambda_1, \dots, \lambda_n) = G_{n,\beta} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^\beta \exp\left(-\frac{\beta}{4} \sum_{i=1}^n \lambda_i^2\right) \prod_i d\lambda_i \quad (1)$$

where $G_{n,\beta}$ is a normalization constant.

For $\beta = 1, 2, 4$, this probability distribution corresponds to the eigenvalues of GOE, GUE and GSE respectively. Other known and related results regarding the β -ensembles will be discussed in the next section. Using the moments method, one can establish the Wigner semicircle law for the β -ensembles, e.g. [4]:

Suppose $\lambda_i, i = 1, \dots, n$ follow the distribution of a β -ensemble. Consider the rescaled eigenvalues $\tilde{\lambda}_i = \frac{\lambda_i}{2\sqrt{n\beta}}$ and the empirical distribution function:

$$N_n(\lambda) = \frac{1}{n} \#\{k : \tilde{\lambda}_k < \lambda\}$$

We have

$$\lim_{n \rightarrow \infty} N_n(\lambda) = \int_{-\infty}^{\lambda} \rho(u) du$$

where

$$\rho(u) = \begin{cases} 0 & \text{for } u > 1 \text{ or } u < -1, \\ \frac{2}{\pi} \sqrt{1-u^2} & \text{for } -1 \leq u \leq 1. \end{cases}$$

Take an r_n -neighborhood O_n of right end of the spectrum $\lambda = 1$, where $r_n = O(n^{-2/3+\epsilon})$ for some $\epsilon > 0$. Normalizing $\tilde{\lambda}_k = 1 - \theta_k r_n$ and

$$\mu_n(\theta_k) = \frac{1}{nr_n^{3/2}}$$

at each θ_k , we obtain a measure μ_n on the real line such that $\mu_n(\mathbb{R}) = r_n^{-3/2}$. The main result of this paper is the proof of the following local semicircle law at the edge of the spectrum:

Theorem 1. *As $n \rightarrow \infty$, μ_n converge weakly in probability on each finite interval to a measure μ concentrated on \mathbb{R}_+ and almost continuous with respect to the Lebesgue measure:*

$$\frac{d\mu}{dx} = \begin{cases} \frac{2\sqrt{2}}{\pi} \sqrt{x} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

If $r_n > n^{\epsilon-2/3}$ for some $\epsilon > 0$, then the measures converge vaguely to μ with probability 1.

Remark 1. *The reason we called the above theorem a local semicircle law is that under the semicircle law, one can show that for a neighborhood of 1 of size of some fixed number δ , the proportion of normalized eigenvalues in that neighborhood converges to the corresponding integral of the semicircle law. The main theorem states that the size of this neighborhood can be much finer, at the scale of $n^{-2/3+\epsilon}$ and the convergence will still hold. In other words, for a neighborhood of size r_n near 1, there are $\Omega nr_n^{3/2}$ eigenvalues. If one rescales the*

total number in this interval by $nr_n^{3/2}$, this number will converge in probability to the appropriate integral of the semicircle law. Note that one would not expect the semicircle shape to hold at the edge when $r_n = n^{-2/3}$. In this sense, the local semicircle law above is almost optimal.

The theorem above will be proven using the moment method, similar to the method by Sinai and Soshnikov [14]:

Theorem 2. *Let A_n be a random matrix whose eigenvalues are given by equation (1) scaled by a factor of $1/2\sqrt{n\beta}$. Let $p_n \rightarrow \infty$ as $n \rightarrow \infty$ and $p_n = O(n^{2/3-\epsilon})$ for some $\epsilon > 0$, then*

$$\mathbb{E}(\text{Tr} A_n^{p_n}) = \begin{cases} 2^{3/2} n (\pi p_n^3)^{-1/2} (1 + o(1)) & \text{if } p_n \text{ is even,} \\ 0, & \text{if } p_n \text{ is odd.} \end{cases}$$

However we shall see in the proof that the combinatorics involved is different and simpler in our case. The derivation of theorem 1 from theorem 2 is the same as in [14] and will be given in the appendix. It should also be remarked that the result in this paper is slightly weaker than that in [14], as in their paper, only $r_n \ll n^{2/3}$ is required.

The structure of the paper is the following. We first introduce a matrix model, proposed by Dumitriu and Edelman [5], for the β -ensembles and recall some related results. We shall then prove theorem 2 in section 3.

For the rest of the paper, we shall use the following notion of an event depending on some index n having overwhelming probability:

Definition 2. *We say an event E holds with overwhelming probability if for all n , $\mathbb{P}(E) \geq 1 - O_C(n^{-C})$ for every constant C .*

It should be observed that a union of n^k events of overwhelming probability for some fixed k still holds with overwhelming probability.

2 Gaussian Beta Ensembles

As mentioned above, the special values of $\beta = 1, 2, 4$ corresponds to eigenvalues of GOE, GUE and GSE respectively and there is an extensive literature on

them. For the edge of the spectrum, Tracy and Widom [18], [19] has been able to establish that, upon rescaling and centering, the top eigenvalue distribution converges to what is known as the Tracy-Widom distribution. Later Soshnikov [15] proved the edge universality for Wigner matrices. However, the case of general β remains unknown until the recent paper by Ramirez, Rider and Virag [13]. One of the obstacles for studying β -ensembles previously is the lack of matrix formulation for them. In their paper [5], Dumitriu and Edelman succeeded in writing down a tridiagonal random matrix model whose eigenvalue distribution is given by equation (1):

Theorem 3 (Dumitriu-Edelman). *Consider the matrix given by*

$$A_{n,\beta} = \begin{pmatrix} N(0,2) & \chi_{(n-1)\beta} & 0 & \dots & \dots & \dots & 0 \\ \chi_{(n-1)\beta} & N(0,2) & \chi_{(n-2)\beta} & 0 & \dots & \dots & 0 \\ 0 & \chi_{(n-2)\beta} & N(0,2) & \chi_{(n-3)\beta} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \chi_{3\beta} & N(0,2) & \chi_{2\beta} & 0 \\ 0 & \dots & \dots & 0 & \chi_{2\beta} & N(0,2) & \chi_{\beta} \\ 0 & \dots & \dots & \dots & 0 & \chi_{\beta} & N(0,2) \end{pmatrix} \quad (2)$$

where $N(0,2)$ denotes a random variable whose distribution follows the Gaussian distribution with mean 0 and variance 2; χ_k is a random variable having a chi distribution with k degree of freedom. The upper triangular part of the matrix consists of independent random variables and the matrix is symmetric. The eigenvalues of this matrix follows the distribution of the β -ensemble.

For GOE, the theorem can be easily proven by successive Householder transformation and this was observed by Trotter [17]. For general β , the idea of the proof is to compute the Jacobian of the diagonalization of the matrix of the form (2).

Another important result is the one by Ramirez, Rider and Virag [13], whose work is based on previous work by Edelman and Sutton [6]. In their paper, Ramirez, Rider and Virag proved that the edge scaling parameter is $n^{2/3}$ and computed the asymptotics of the rescaled limiting distribution of the β -ensembles:

Theorem 4 (Ramírez, Rider, Virag). *Let $\lambda_1, \dots, \lambda_k$ be the k largest eigenvalue of the β -ensemble, in descending order, and*

$$\mathcal{H}_\beta = -\frac{d^2}{dx^2} + x + \frac{2}{\beta}b'_x$$

where b' denote a white noise. Let $\Lambda_1, \dots, \Lambda_k$ be the k lowest eigenvalues of \mathcal{H}_β , then

$$n^{1/6}(\lambda_n - \sqrt{2n\beta})$$

converges in distribution to $(\Lambda_1, \dots, \Lambda_k)$ Moreover the asymptotics of Λ_1 are given by:

$$\mathbb{P}(\Lambda_1 > a) = \exp\left(-\frac{2}{3}\beta a^{3/2}(1 + o(1))\right)$$

and

$$\mathbb{P}(\Lambda_1 < -a) = \exp\left(-\frac{1}{24}\beta a^3(1 + o(1))\right)$$

The proof of the above theorem is based on the observation that the tridiagonal matrix model can be viewed as a discretization of certain stochastic differential operator \mathcal{H}_β and the probability distribution of the largest eigenvalue of the tridiagonal matrix will converge to the probability distribution of the eigenvalues associated stochastic differential operator as $n \rightarrow \infty$. This point process is given by the so-called *Airy* $_\beta$ process and upon rescaling will converge to μ in our main theorem. Notice that if we consider the operator \mathcal{H}_∞ , which is the free Laplacian, then the spectral measure of the operator under proper rescaling is μ .

Of course, these two results are only the tip of the iceberg of the vast field of random matrices, there are excellent monographs, e.g. [1], [10] that discuss the latest developments in details.

We see from the above theorem that, at the scaling of $n^{2/3}$, the limiting distributions are different for each β at the edge of the spectrum. On the other hand, the Wigner semicircle law is universal for all β , one can therefore ask the question at which scale does the edge look different for different β , which leads us to consider the problem at hand. Moreover, one can also ask, given the availability of a matrix representation, whether the classical moment method will be able to give us information at the edge of the spectrum. It is also part of the goal of this paper to explore this possibility. Lastly, a local semicircle law is always of interest. The power of a local semicircle law can be seen in recent papers of Erdős, Schlein and Yau [7], [8] and in Tao and Vu [16] in their proof of universality in Wigner ensembles and also in the recent paper of Bourgade, Erdős and Yau [3]. Bao and Su [2] has also established through the Brownian Carousel a local semicircle law in the bulk for the Gaussian β -ensemble. Also, one does not expect to prove a local result at the bulk using the moments as higher moments highlight the eigenvalues at the edge.

3 Combinatorial considerations

Our goal is to understand the trace of higher degrees of the matrix $A_{n,\beta}$. For simplicity of notations, we consider the matrix scaled by $1/\sqrt{\beta}$, so that the off diagonal entries have variance $n, n-1, \dots, 1$ and the diagonal entries have variance $2/\beta$. The only difference is that we have to normalize the resulting calculations of moments by $2^{p_n} n^{p_n/2}$. The traces are given by

$$\mathbb{E}(\text{Tr} A_{n,\beta}^k) = \mathbb{E} \sum_{\mathcal{P}} \xi_{i_0 i_1} \xi_{i_1 i_2} \dots \xi_{i_{k-1} i_0} \quad (3)$$

where \mathcal{P} is the set of indices $\{i_0, i_1, \dots, i_{k-1}\}$

It is customary to encode this information as a graph and to study the combinatorics associated with this graph:

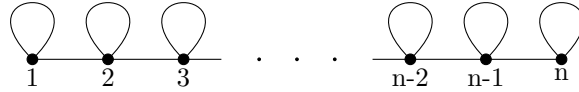


Figure 1: Graph representation of a tridiagonal matrix

We first note an easy simplification of our problem at hand, due to Dumitriu (see [4] for the proof):

Lemma 1. $\mathbb{E}(\text{Tr} A_n^k) = n \mathbb{E}(A_n^k)_{11}$

The plan now is to first calculate the contributions from paths that do not go through self loops, i.e. paths where $i_j \neq i_{j+1}$ for all j ; after that we shall prove that the other paths contribute asymptotically lower order terms to the sum.

Let us discuss of the odd moments first. Even though the offdiagonal entries are not symmetric distributions and have nonzero odd moments. One notices that the coefficients of the characteristics equation for the matrix involves only even powers of the offdiagonal entries, so we can replace the χ -distribution by a symmetric distribution that has the same even moments as the χ -distribution and the joint distribution of eigenvalues of the resulting matrix will be the same as that of the original. Since the distribution of the trace depends only on that of the eigenvalues, we see that the odd moments vanish. If $f(x)$ is the density of the χ -distribution, then the distribution $g(x)$ equals to $f(x)/2$ for $x > 0$ and $f(-x)/2$ for $x < 0$ will be the desirable replacement.

We now return to the calculations of the even moments. Associate with each path on the graph a Dyck path of length $2k$, where travelling to the right corresponds to an up edge and travelling to the left corresponds to a down edge. Throughout the rest of the paper, we shall most of the time think of Dyck paths as probabilistic objects, i.e. samples from a random walk X of length $2k$ conditioned on $X_0 = X_{2k} = 0$ and $X_m \geq 0$ for $0 \leq m \leq 2k$.

Before we start the proof in earnest, let us first take a step back and recall some relevant known facts and see what could cause problems:

1) The number of Dyck paths of length $2k$ is given by the k -th Catalan number C_k . 2) The $2l$ -th moment of a chi-distributed random variable of degree with freedom k is given by $\prod_{i=0}^l (k + 2i)$.

Each path in the sum (3) will contribute $\prod_{i=0}^k (n + j_i)$ for some j_i . The hope is that asymptotically, for $k = o(n^{2/3})$, the product will be $n^k(1 + o(1))$ and the worry is that some of the j_i gets too small (negative) or too big. This corresponds to the case where the path goes too far to the right or stay on the same edge on the graph for too many steps. Therefore our goal is to show that there are very few of these problematic paths that they do not affect the total sum as n goes to infinity.

The first step is to connect Dyck paths with a true random walk, so that we can use many of the known results and nice properties of the latter. We start with the following well-known lemma, which establishes the relationship between Dyck paths and Bernoulli bridges.

Lemma 2. *There is a one-to-one correspondance between $\{1, \dots, k + 1\} \times \{\text{Dyck paths of length } 2k\} \leftrightarrow \{\text{Bernoulli Bridges of length } 2k\}$*

Proof. The proof is to provide an explicit map. Attach a down path to the end of a Dyck path, resulting in a total of $k + 1$ down paths and k up paths. Pick one of the $k + 1$ down paths and make a cut before the down path. The Bernoulli bridge (from $(0, 0)$ to $(2k, -1)$ with first path always downwards) is constructed by moving the second half to the origin and attaching the first half to it;

□

The next lemma, relating the probability of an event of a Bernoulli bridge with one of a random walk, is first proven by Khorunzhiy and Marckert [11]:

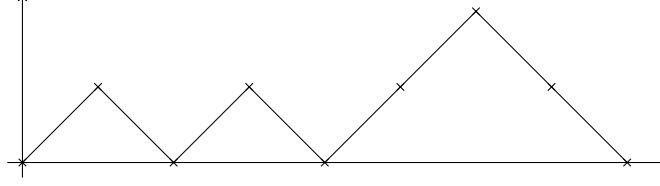


Figure 2: A sample Dyck path

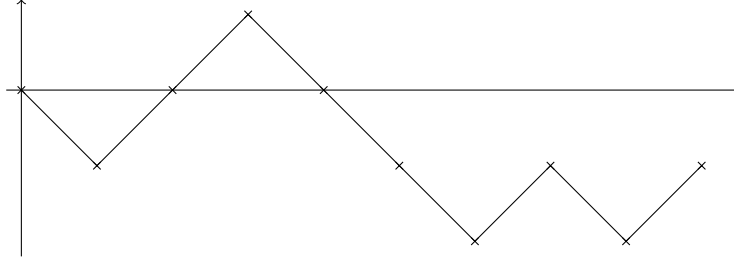


Figure 3: The Bernoulli bridge corresponding to the above Dyck path cutting at the second down edge

Lemma 3 (Khorunzhiy, Marckert). *Let $S_k = X_1 + \dots + X_k$ be a simple random walk. Let \mathcal{F}_l be the measurable events generated by X_1, \dots, X_l and let \mathbb{P}_l^w be the probability measures induced on the paths by X_1, \dots, X_l . Let \mathbb{P}_l^b be the measure \mathbb{P}_l^w conditioned on $X_0 = 0$ and $S_l = -1$. Then for any \mathcal{F}_k -measurable event A_k ,*

$$\mathbb{P}_{2k+1}^b(A_k) \leq C \mathbb{P}_{2k+1}^w(A_k) = C \mathbb{P}_k^w(A_k) \quad (4)$$

for some constant C independent of n .

Proof. The second equality is obvious as A_k is \mathcal{F}_k -measurable. Suppose A is \mathcal{F}_k -measurable, then

$$\mathbb{P}_{2k+1}^w(A|S_k = m, S_{2k+1} = -1) = \mathbb{P}_{2k+1}^w(A|S_k = m) = \mathbb{P}_k^w(A|S_k = m),$$

thus,

$$\begin{aligned} \mathbb{P}_{2k+1}^b(A) &= \mathbb{P}_{2k+1}^w(A|S_{2k+1} = -1) \\ &= \sum_m \mathbb{P}_k^w(A|S_k = m) \mathbb{P}_{2k+1}^w(S_k = m|S_{2k+1} = -1). \end{aligned}$$

On the other hand,

$$\begin{aligned}\mathbb{P}_{2k+1}^w(S_k = m | S_{2k+1} = -1) &= \frac{\binom{k}{m} \binom{k+1}{m+1}}{\binom{2k+1}{m+1}} \\ &= \frac{\binom{k}{m}}{2^k} \frac{2^k \binom{k+1}{m+1}}{\binom{2k+1}{m+1}} \\ &\leq \mathbb{P}_k^w(S_k = m) C_0,\end{aligned}$$

where $C_0 := \sup_k \sup_m \frac{2^k \binom{k+1}{m+1}}{\binom{2k+1}{m+1}}$ is finite. It follows that

$$\begin{aligned}\mathbb{P}_{2k+1}^b(A) &\leq C_0 \sum_m \mathbb{P}_k^w(A | S_k = m) \mathbb{P}_k(S_k = m) \\ &= C_0 \mathbb{P}_k^w(A)\end{aligned}$$

□

As discussed earlier, we would like to show that the probability of crossing too far to the right of the graph is small and this is indeed the case:

Proposition 1. *Let S_{2k} be a random walk conditioned on $S_0 = S_{2k} = 0$ and $S_l \geq 0$ for all $0 \leq l \leq 2k$. Then for all k*

$$\mathbb{P}\left(\max_{0 \leq l \leq 2k} S_k \geq \lambda k^{1/2}\right) \leq C \exp(-c\lambda^2) \quad (5)$$

for some constants c and C .

Proof. Using the notations from the previous lemma, denote the Bernoulli excursion of length $2k$, the Bernoulli bridge from $(0, 0)$ to $(2k, 0)$, and the random walk of length $2k$ by S_{2k}^e , S_{2k}^b , S_{2k}^w respectively.

By lemma 2, we have

$$\max_{0 \leq l \leq 2k} S_l^b - \min_{0 \leq l \leq 2k} S_l^b - 1 \leq \max_{0 \leq l \leq 2k} s_l^e \leq \max_{0 \leq l \leq 2k} S_l^b - \min_{0 \leq l \leq 2k} S_l^b + 1$$

So it suffices to show that

$$\mathbb{P}\left(\max_{0 \leq l \leq 2k} S_k^b - \min_{0 \leq l \leq 2k} S_k^b \geq \lambda k^{1/2}\right) \leq C \exp(-c\lambda^2)$$

Moreover, if

$$\max_{0 \leq l \leq 2k} S_k^b - \min_{0 \leq l \leq 2k} S_k^b \geq \lambda k^{1/2},$$

then either

$$\max_{0 \leq l \leq k} S_k^b - \min_{0 \leq l \leq k} S_k^b \geq \frac{1}{2} \lambda k^{1/2} \quad (6)$$

or

$$\max_{k+1 \leq l \leq 2k} S_k^b - \min_{k+1 \leq l \leq 2k} S_k^b \geq \frac{1}{2} \lambda k^{1/2}.$$

By reversing the Bernoulli bridge, it suffices to show the first inequality. The event satisfying the inequality (6) is contained in the union of the events

$$\max_{0 \leq l \leq 2k} S_k^b \geq \frac{\lambda}{4} k^{1/2} \text{ and } \min_{0 \leq l \leq 2k} S_k^b \leq \frac{-\lambda}{4} k^{1/2}$$

By lemma 3,

$$\begin{aligned} \mathbb{P}(\max_{0 \leq l \leq k} S_k^b \geq \frac{-\lambda}{4} k^{1/2}) &\leq C_0 \mathbb{P}(\max_{0 \leq l \leq k} S_k^w \geq \frac{-\lambda}{4} k^{1/2}) \\ &\leq 2C_0 \mathbb{P}(S_k \geq \frac{-\lambda}{4} k^{1/2}) \end{aligned} \quad (7)$$

where the second inequality is implied by the reflection principle. Hoeffding's inequality then implies the desired inequality:

$$\mathbb{P}(\max_{0 \leq l \leq k} S_k^b \geq \frac{-\lambda}{4} k^{1/2}) \leq C \exp(-c\lambda^2)$$

□

Next, we shall consider the potential problem of a path staying on the same edge too many times. Again, the probability of this event is very small:

Proposition 2. *Let $S_l, 0 \leq l \leq 2k$ be the Bernoulli excursion as above. Let T_i denote the number of l such that $S_l = i$, for $i = 1, \dots, 2k$, then*

$$\mathbb{P}(\max_i T_i \geq k^{1/2+\epsilon}) \leq C \exp(-ck^{2\epsilon'})$$

for any $\epsilon > \epsilon' > 0$ and some constants C, c independent of k, λ .

Proof. As in the previous proposition, we shall connect the quantity in question with a similar quantity of a random walk. Clearly, since by mapping Dyck paths to Bernoulli bridges, the times one visits a level is mapped to at most two different levels, we have

$$\mathbb{P}(\max_i T_i^e \geq \lambda k^{1/2}) \leq \mathbb{P}(2 \max_i T_i^b \geq \lambda k^{1/2})$$

and as before we split the bridge into two halves

$$\mathbb{P}(2 \max_i T_i^b \geq \lambda k^{1/2}) \leq \mathbb{P}(\max_{0 \leq i \leq k} T_i^b \geq \frac{\lambda}{4} k^{1/2}) + \mathbb{P}(\max_{k+1 \leq i \leq 2k} T_i^b \geq \frac{\lambda}{4} k^{1/2})$$

By lemma 3, it suffices to show that

$$\mathbb{P}(\max_{0 \leq i \leq k} T_i^w \geq \frac{\lambda}{4} k^{1/2}) \leq C \exp(-c\lambda^2)$$

We will in fact first show that, for each $-k \leq 0 \leq k$,

$$\mathbb{P}(T_i^w \geq \frac{\lambda}{4} k^{1/2}) \leq C \exp(-c\lambda^2)$$

Let $q_{2k,r}$ be the probability that a random walk of length $2k$ returns to 0 exactly r times. It is well known (e.g. [12]) that

$$q_{2k,r} = \frac{1}{2^{2k-r}} \binom{2k-r}{k}$$

We can then approximate the probability that there are less than $k^{1/2+\epsilon}$ returns to 0 for a random walk of length $2k$ using Stirling's formula

$$\mathbb{P}(\text{there are less than } k^{1/2+\epsilon} \text{ returns to 0}) \sim \sqrt{\frac{2}{\pi}} \int_0^{k^{1/2+\epsilon}} e^{-t^2/2} dt$$

For levels other than 0, we can condition on the first hitting time and similarly show that

$$\mathbb{P}(T_i \geq k^{1/2+\epsilon}) \leq C \exp(-ck^{2\epsilon})$$

Therefore, we have the bound

$$\mathbb{P}(\max_i T_i \geq k^{1/2+\epsilon}) \leq C \exp(-ck^{2\epsilon'})$$

for any $\epsilon' < \epsilon$

□

With the two propositions above, we are ready to compute the contribution by the main paths. First of all the number of Dyck paths of length $2k$ is given by the Catalan number $C_k = \frac{1}{k+1} \binom{2k}{k}$. We have shown that $\frac{1}{k+1} \binom{2k}{k} (1 - o(1))$ of these paths have 'good properties', i.e. they are not in the exceptional set in either propositions. The contribution of each paths is of order $n^k (1 + o(1))$ when $k = o(n^{2/3})$. To see this, we simply compute the contribution

$$(n + i_1) \dots (n + i_k)$$

for that particular path. Not being in the exceptional set in the first proposition means that $i_j \geq -n^{1/3-\delta}$ for any $\delta > 0$ and not being in the exceptional set in the second proposition means that $i_j \leq n^{1/3-\delta}$ for any $\delta > 0$. Therefore $\sum_j i_j \leq n^{1-\delta}$ and the claim follows.

It remains to show that the other paths do not contribute asymptotically to the sum. Let us first tackle the Dyck paths that do not have 'good properties'.

Consider the paths where $\max_i T_i$ is of the order q_n where $q_n = p_n^{1/2+l}$ for some $l > 0$, each of the paths will contribute $O(n^{p_n}(1 + \exp(n^l)))$. By the previous proposition, the probability density of these paths are asymptotically of the order $O(\exp(-n^{2l-\delta}))$. Integrating over these paths, we can conclude that they are asymptotically of $o(n^{p_n})$.

Lastly we compute the paths where there are at least one loop, i.e., on the path $\{1, i_1, i_2, \dots, i_{p_n-1}, 1\}$, there exists l such that $i_l = i_{l+1}$:

Proposition 3. *The contribution to the sum from paths with $2k$ loops, i.e. paths $\{1, i_1, i_2, \dots, i_{p_n-1}, 1\}$ where there exists l_1, \dots, l_{2k} such that $i_{l_j} = i_{l_j+1}$ for $j = 1, \dots, 2k$ is of $O(C_{p_n/2} n^{\frac{p_n}{2}-k\epsilon})$*

Proof. First notice that the number of loops must be even as the odds moments of a Gaussian random variable are zero. As in proposition (2), we let T_i denote the number of times the path stays on level i . The number of paths with $2k$ loops are bounded by

$$\frac{1}{\frac{p_n}{2} - k + 1} \binom{p_n - 2k}{\frac{p_n}{2} - k} \sum_{2a_1, \dots, 2a_l: \sum_1^l 2a_i = 2k} \binom{T_i}{2a_i}, \quad (8)$$

since the loops have to be paired on the 'same level' and removing the loops give a Dyck path of length $p_n - 2k$. With overwhelming probability, each path contributes $O(n^{\frac{p_n}{2}-k} \prod_i (2a_i - 1)!!)$. Indeed, if we remove the $2k$ loops, it reduces to a Dyck path that the earlier analysis applies, hence the factor $n^{\frac{p_n}{2}-k}$ and each loop on the same level contributes the $2a_i$ -th moment of the Gaussian random variable, which is of the order $(2a_i - 1)!!$. For a given k , the contribution from the terms in equation (8) is dominated by the term where $l = k$ and $a_1 = \dots = a_l = 2$ (one way to see this is that this is a similar argument as the one which shows that trace of moments of Wigner matrices are dominated by paths that traverse each edge twice after considering only the contribution from the loops, see e.g. [14], the key here being the moments of Gaussian does not grow too fast; another way is to calculate the contribution when $l < k$. This is bounded by $\binom{n^{1/3-\epsilon}}{l} l^{k-l} n^{2k/3-k\epsilon} (\lceil \frac{k}{l} \rceil!!)^l$, the binomial term coming from picking l levels for the loops, the second term being each level must be filled by at least two loops, and the remaining loops can be anywhere, the third term is just the maximum possible of $\prod_i \binom{T_i}{a_i} (2a_i - 1)!!$ and as explained below, $\max_i T_i < n^{1/3-\epsilon}$. Summing over all $l < k$ yields the desired dominance.) and is bounded by $n^{k/3-k\epsilon} n^{2k/3-k\epsilon}$ for $p_n = O(n^{2/3-\epsilon})$, since we have $\max_i T_i = O(n^{1/3-\epsilon})$ with overwhelming probability and therefore $\binom{T_i}{a_i} = O(n^{a_i/3-\epsilon})$ with overwhelming probability and moreover, by proposition (1), the number of different T_i to choose from in the sum is of $O(n^{1/3-\epsilon})$ with overwhelming probability and so the number of choice is bounded by $n^{k/3-k\epsilon}$. The exceptional paths can be dealt with similar to the ones without loops. Using the trivial bound of $C_{p_n/2-k} < C_{p_n}$ result in the proposition. \square

If we sum over all $k \geq 2$ in the proposition above, we found that the contribution from the paths with loops is of the order $O(C_{p_n/2} n^{p_n/2-\epsilon})$.

Combining all of the above, we have that $\mathbb{E}(A_{n\beta}^{p_n})_{11}$ is given by

$$C_{p_n/2} n^{p_n/2} (1 + o(1))$$

when p_n is even and 0 when p_n is odd. Here C_k denotes the k -th Catalan number, which counts the number of Dyck paths between 0 and $2k$. It is well known that:

$$C_k \sim \frac{4^k}{k^{3/2} \sqrt{\pi}}$$

and the result follows from lemma 1 after we rescaled by $2^{p_n} n^{p_n/2}$.

4 Conclusion

In this article we prove a local semicircle law at the edge of the spectrum for β -ensembles. The proof relies on calculations of moments and can be applied to other tridiagonal matrices satisfying fairly mild conditions: that the diagonal entries remain $O(1)$ and the growth of the moments of the off diagonal entries is not too fast. It would be interesting to consider the same question for the β -Laguerre ensembles.

5 Acknowledgements

The author would like to thank his advisor Prof. Ya. G. Sinai for his suggestion of the problem and his continual support and Prof. Alice Guillonet for providing useful remarks and discussions. Finally, the author would like to thank the referee for pointing out various typos and making a number of suggestions to improve the article.

6 Appendix

We shall provide the derivation of theorem 1 from theorem 2 for the readers' convenience.

Proof. Proof of theorem 1 To prove the theorem, it suffices to establish the convergence of the Laplace transform of the measures μ_n to that of μ (see e.g. [9] for the relationship between convergence of Laplace transform and convergence in distribution):

$$\int_{-\infty}^{\infty} e^{-c\theta} d\mu_n(\theta) = \frac{1}{nr_n^{3/2}} \sum_{k=1}^n e^{-c\theta_k} \xrightarrow{n \rightarrow \infty} \int_0^{\infty} e^{-c\theta} \frac{2\sqrt{2}}{\pi} \sqrt{\theta} d\theta = \sqrt{\frac{2}{\pi c^3}}$$

Let $s_n = cr_n^{-1}/2$ and $p_n = 2s_n$. We claim that the sequence $n^{-1}r_n^{3/2}\text{Tr}A^{2s_n}$ converges in probability to

$$\lim_{n \rightarrow \infty} \frac{1}{nr_n^{3/2}} \mathbb{E}(\text{Tr}A^{2s_n}) = \lim_{n \rightarrow \infty} \frac{1}{nr_n^{3/2}} \frac{n}{\sqrt{\pi s_n^3}} = \frac{2\sqrt{2}}{\sqrt{\pi c^3}}$$

This follows from Chebyshev's inequality once we show that the variance of $n^{-1}r_n^{3/2}\text{Tr}A^{2s_n}$ is of $o(1)$. This is true as:

$$\begin{aligned} \text{var}(n^{-1}r_n^{3/2}\text{Tr}A^{2s_n}) &= n^{-2}r_n^{-3} \text{var}(\text{Tr}A^{2s_n}) \\ &\leq n^{-1}r_n^{-3} \mathbb{E}((\text{Tr}A^{2s_n})^2) \leq 2n^{-1}r_n^{-3} \mathbb{E}(\text{Tr}A^{4s_n}) \end{aligned}$$

By theorem 2, the last quantity is of $o(1)$. Similarly, $n^{-1}r_n^{-3/2}\text{Tr}A^{2s_n+1}$ converges in probability to zero. Let

$$\lambda_k = 1 - \theta_k r_n \lambda_k \geq 0 \quad (9)$$

$$\lambda_j = -1 + \tau_j r_n \lambda_j \leq 0 \quad (10)$$

Then by definition, we have

$$\text{Tr}A^{2s_n} = \sum_k (1 - r_n \theta_k)^{2((c/2)r_n^{-1})} + \sum_j (1 - \tau_j r_n)^{2((c/2)r_n^{-1})} \quad (11)$$

and

$$\text{Tr}A^{2s_n+1} = \sum_k (1 - r_n \theta_k)^{2((c/2)r_n^{-1})+1} - \sum_j (1 - \tau_j r_n)^{2((c/2)r_n^{-1})+1} \quad (12)$$

If $|\theta_k| \leq r_n^{-1/3}$ and $|\tau_j| \leq r_n^{-1/3}$, the corresponding terms in (11) and (12) are $e^{-c\theta_k}(1 + o(r_n^{1/3}))$ and $e^{-c\tau_j}(1 + o(r_n^{1/3}))$. On the other hand, using the estimate of the expectation of $\text{Tr}A^{4s_n}$, the subsums in (11) and (12) over θ_k and τ_j such that $|\theta_k| > r_n^{1/3}$ and $|\tau_j| > r_n^{1/3}$ converge to zero in probability. The same holds for the subsums over these θ_k and τ_j of $\sum_k e^{-c\theta_k}$ and $\sum_j e^{-c\tau_j}$. Therefore, we have

$$(\text{Tr}A^{2s_n} + \text{Tr}A^{2s_n+1} - 2 \sum_k e^{-c\theta_k}) \frac{1}{nr_n^{3/2}} \rightarrow 0$$

and

$$\frac{1}{nr_n^{3/2}} \sum_k e^{-c\theta_k} \rightarrow \sqrt{\frac{2}{\pi c^3}}$$

If $r_n = n^{-\gamma}$ where $\gamma < 2/3$, then it follows from theorem 2 that the normalized traces $n^{-1}r_n^{-3/2}\text{Tr}A^{2s_n}$ and $n^{-1}r_n^{-3/2}\text{Tr}A^{2s_n+1}$ converges to nonrandom limits with probability 1, and hence the convergence of the measures μ_n in the main theorem also occurs with probability 1. \square

References

- [1] Anderson, G.W., Guionnet, A. and Zeitouni, O.: An Introduction to Random Matrices, *Cambridge Univ. Press* (2010)
- [2] Bao, Z., and Su, Z.: Local semicircle law and Gaussian fluctuation for Hermite β -ensemble. (2011) Preprint, arXiv:math.PR/1104.3431
- [3] Bourgade, P., Erdős, L., and Yau, H.-T.: Universality of General β -ensemble. (2011) Preprint, arXiv:math.PR/1104.2272
- [4] Dumitriu, I.: Eigenvalue statistics for beta ensembles, *Ph.D. thesis, MIT* (2003)
- [5] Dumitriu, I. and Edelman, A.: Matrix models for beta ensembles, *J Math. Phys.* **43** (11), 5830-5847 (2002)
- [6] Edelman, A. and Sutton, B.: From random matrices to stochastic operators *J. Stat. Phys.* **127** (6), 1121-1165 (2007)
- [7] Erdős, L., Schlein, B. and Yau, H.-T.: Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices *Ann. Probab.* **37**(3), 815-852 (2008)
- [8] Erdős, L., Schlein, B., and Yau, H.-T.: Local semicircle law and complete delocalization for Wigner random matrices. *Commun. Math. Phys.* **287**, 641-655 (2009)
- [9] Feller, W., An introduction to Probability Theory and its Applications, Vol. 2 *Wiley and Sons* (1971)
- [10] Forrester, P.: Log-gases and Random Matrices, *Princeton Univ. Press* (2010)
- [11] Khorunzhiy, O. and Marckert, J.-F.: Uniform bounds for exponential moment of maximum of a dyck path, *Elect. Comm. in Probab.* **14** 327-333 (2009)
- [12] Revesz, P.: Random walk in random and non-random environments, *World Scientific* (2005)

- [13] Ramírez, J., Rider, B. and Virag, B: Beta ensembles, stochastic Airy spectrum and a diffusion. (2007) Preprint, arXiv:math.PR/0607331.
- [14] Sinai, Ya. G. and Soshnikov, A.B.: A refinement of Wigner's semicircle law in a neighborhood of the spectrum edge for random symmetric matrices, *Funct. Anal. and Its Appl.* **32**(2) (1998)
- [15] Soshnikov, A.B.: Universality at the edge of the spectrum in Wigner random matrices, *Comm. Math. Phys.* **207** 697-733 (1999)
- [16] Tao, T. and Vu, V.: Random matrices: Universality of local eigenvalue statistics, (2009) Preprint, arXiv:0906.0510v10
- [17] Trotter, H.: Eigenvalue distributions of large Hermitian matrices; Wigner's semicircle law and a theorem of Kac, Murdoch and Szego. *Adv. in Math.* **54**(1) 67-82 (1984)
- [18] Tracy, C., and Widom, H.: Level spacing distributions and the Airy kernel. *Comm. Math. Phys.* **159**(1) 151-174 (1994)
- [19] Tracy, C., and Widom, H.: On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.* **177**(3) 727-754 (1996)